



Combating Misinformation by Sharing the Truth: a Study on the Spread of Fact-Checks on Social Media

Jiexun Li¹ · Xiaohui Chang²

Accepted: 19 May 2022 / Published online: 11 June 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Misinformation on social media has become a horrendous problem in our society. Fact-checks on information often fall behind the diffusion of misinformation, which can lead to negative impacts on society. This research studies how different factors may affect the spread of fact-checks over the internet. We collected a dataset of fact-checks in a six-month period and analyzed how they spread on Twitter. The spread of fact-checks is measured by the total retweet count. The factors/variables include the truthfulness rating, topic of information, source credibility, etc. The research identifies truthfulness rating as a significant factor: conclusive fact-checks (either true or false) tend to be shared more than others. In addition, the source credibility, political leaning, and the sharing count also affect the spread of fact-checks. The findings of this research provide practical insights into accelerating the spread of the truth in the battle against misinformation online.

Keywords Social media · Information sharing · Misinformation · Credibility · Fact-check · Data analysis

1 Introduction

Misinformation has long existed in human history (Olan et al., 2022). While the internet becomes the most popular information hub, the creation and dissemination of misinformation have generally grown on the internet, as well. Furthermore, on social media sites, social features such as sharing, liking, and following have made these platforms a perfect hotbed for misinformation. Misinformation can be defined as any information that is inaccurate, wrong, or false, regardless of whether there is an intent to mislead. Disinformation, in contrast, involves deliberately fabricated information or manipulated narratives with the intention of propaganda and/or harm (Pennycook & Rand, 2021; Rodrigo et al., 2022). The rapid growth of misinformation over the internet can cause various types of harms, including life, injury, income, business, emotion, trust, reputation, discrimination, connection, isolation, safety, access, privacy, decision, and confusion harms (Tran et al., 2021). Such negative

impacts are particularly substantial and concerning during natural, humanitarian, or political crises. For example, during the COVID-19 pandemic since early 2020, many people have encountered inaccurate or fabricated information, ranging from the origins (e.g., 5G signal), medicine (e.g., Vitamin-C, hydroxychloroquine, Clorox to treat COVID-19), to vaccines (e.g., Bill Gates to implant microchips in people's bodies via vaccines). These COVID-related fake news and conspiracy theories create confusion, manipulate people's behaviors, and undermine the credibility of science (Hopf et al., 2019; Roozenbeek et al., 2020; Tasnim et al., 2020). Furthermore, fake news and disinformation are often fabricated for political attacks, which can ignite hatred, cause ideological polarization, and lead to social instability and compromised democracies (Au et al., 2021; Yusof et al., 2020). During the 2020 US Presidential Election, a sheer amount of misinformation emerged. Even the former U.S. President Donald J. Trump has been accused of sharing misinformation about the election results and incitement of violence during the U.S. Capitol Hill Riot in January 2021 (Olan et al., 2022). Trump's accounts were then suspended or terminated by several social media platforms.

To combat misinformation on social media, the first step is to detect it. This is a difficult task because misinformation is often intentionally fabricated to mislead people. Many researchers have developed techniques to automatically

✉ Jiexun Li
Jiexun.Li@wwu.edu

¹ Western Washington University, Bellingham, WA 98225, USA

² Oregon State University, Corvallis, OR 97331, USA

detect misinformation (Kumar & Shah, 2018). In many cases, text features alone do not provide sufficient traits to determine if the information is true or false. Other information such as context, user engagement, and social behaviors may be key signals for misinformation detection (Kumar & Shah, 2018; Shu et al., 2019). To mitigate the negative impacts of misinformation, organizations (e.g., PolitiFact and Snope) have emerged as fact-checkers that seek to verify information and detect fake news. The implementation of fact-checking tools is critical in building acceptance and trust in society (Olan et al., 2022).

The fact-checking results validated by professionals will not make a difference unless they are disseminated on the internet wider and faster than the fake news. This leads to the second step in combatting misinformation: spread the truth. This is particularly challenging in a politically divided atmosphere, where people tend to tune into like-minded sources (e.g., political leaders, influencers, and news media). A Pew Research Center study (Pewresearch, 2020) shows that 64% of Americans say social media has a mostly negative effect on the way things are going in the U.S. today. Many criticize social media's role in the manifestation of confirmation bias (Modgil et al., 2021) and the creation of echo chambers, in which (mis)beliefs are reinforced via communication and repetition of similar ideologies from like-minded peers or sources (Vicario et al., 2016). This tends to exacerbate polarization and hinder the spread of the truth.

Combating misinformation requires not only consistent fact-checking but also adequate propagation of the fact checks. While many prior works have accomplished successes in the first step, the second step of spreading the truth is often omitted and calls for more in-depth research. Hence, it is of critical importance to investigate how users react to fact-checks and disseminate them on the internet. Many factors may affect information diffusion on the internet. In this research, we aim to answer the following questions: *What are the main factors that can affect the spread of fact-checks on social media? How can we promote the spread of fact-checks to mitigate the damages caused by misinformation?* Based on information processing and information diffusion theories, our research examines factors associated with not only the target information fact-checked but also the sources and how the fact-checks are published. Our findings can provide practical insights into how to quickly spread the truth in the battle against misinformation on the internet.

The remainder of the paper is organized as follows. In Section 2, we review related work on misinformation detection and fact-checking. In Section 3, we describe our data collection and methodology of examining different factors' effects on the spread of information. Next, we demonstrate and discuss the results in Section 4. Finally, Section 5

concludes this paper with implications of our findings and future research directions.

2 Related Work

Facing the increasing amount of misinformation spread over the internet, many researchers have attempted to mitigate its negative impact by studying this phenomenon from the following two aspects: (1) *detection of misinformation* and (2) *spread of misinformation*.

2.1 Detection of Misinformation

The first step to fight misinformation is to detect it. In existing works, the detection of misinformation is often formulated as a classification problem that tries to determine if a message is true or false based on various features (Kumar & Geethakumari, 2014; Shu et al., 2017; Wu et al., 2019). Most works have focused on one of the two types of misinformation: fact-based (e.g., fake news) and opinion-based (e.g., fake reviews) (Kumar & Shah, 2018). Particularly, disinformation is often created in ways that intentionally make it more captivating and believable to readers. Misinformation detection involves various features, including not only the text itself but also how it is presented, by whom, and in what format and context (Sloan et al., 2017; Kumar & Shah, 2018) reviewed the different characteristics of misinformation. They found that opinion-based misinformation often exhibits characteristics such as duplications, short lengths, over-exaggeration, skewed rating distribution, and short inter-arrival times, while fact-based misinformation often tends to be longer, generates more confusion, and is created by newer accounts that are tightly connected. Luca and Zervas (2016) built empirical data models to investigate the economic incentives to commit review fraud. They identified that business type, performance, and competition are among the main factors for businesses in committing fake review spam. Shu et al. (2019) also demonstrate the predictive power of social contexts for fake news detection. Hence, domain-specific features of information sources and contexts are critical to the success of misinformation detection.

Although these automated learning models may have demonstrated some promises, we are not yet ready to trust them alone, considering the large volume and variety of information in terms of topics, contexts, and sources. Accurate detection of misinformation still largely relies on a lot of manual effort for in-depth investigation and analysis. Allcott and Gentzkow (2017) attribute the spread of misinformation to the lack of “*third-party filtering, fact-checking, or editorial judgment*” on the internet. Recognizing the significance of this problem, many organizations (e.g., PolitiFact, Snope) and big-tech companies (e.g., Google and Facebook) have

stepped up as fact-checkers for the public to verify information spread on the internet. These fact-checkers mainly focus on investigating news stories, rumors, or statements made by political figures/organizations based on in-depth investigation and analysis. They publish their judgments by posting on their website, distributing on social media platforms, or taking direct action (e.g., deleting or flagging) against misinformation. In addition, rather than examining information itself, Pennycook and Rand (2019) looked into the credibility of information sources. Based on their finding that laypeople are often capable of differentiating news source quality, they suggested incorporating crowdsourced judgments into ranking algorithms to fight misinformation on social media.

2.2 Spread of Misinformation

On social media, individuals build relationships with each other in various forms, including “follow” (Rabelo et al., 2012a, b; Speriosu et al., 2011; Tan et al., 2011), “mention” (Conover et al., 2011; Tan et al., 2011), or “retweet” (Conover et al., 2011; Rajadesingan & Liu, 2014; Wong et al., 2013). These relationships further form a social network that enables the diffusion of information. Research on social media analytics confirms the power of “*homophily*” (Lazarsfeld & Merton, 1954), i.e., a phenomenon of “birds of a feather flock together” (McPherson et al., 2001). Users who are “connected” by a mutual relationship are more likely to share common ideologies/opinions.

On social media platforms like Twitter, retweeting is one of the easiest ways to share and spread information. Compared to formulating an original tweet by oneself, retweeting costs little to construct a message. Many researchers have investigated the underlying motivations of Twitter users for retweeting. Lee et al. (2015) attributed retweeting to users’ prosocial motivation of contributing to their community (Dovidio, 1984) from three dimensions: egoistic, altruistic, and reciprocity. Egoistic motivation sees people as self-oriented individuals and their prosocial behaviors are ultimately self-serving (Carlo et al., 1991). By contrast, altruistic motivation attributes prosocial behaviors to genuine concern and empathy for others’ welfare (Cialdini et al., 1987). Others consider reciprocity function as a fundamental interactive principle in online communities. Users share information because they believe that participation and interaction can positively build up their communities (Wasko & Faraj, 2000). Retweeting can also serve as a conversational practice for communication (Boyd et al., 2010), maintaining social relationships (Recuero et al., 2011), self-expression (Lee et al., 2015), obtaining/updating information (Hwang & Shim, 2010), seeking feedback (Abdullah et al., 2017), etc. While homophily is a key factor that drives retweeting, it is also evident that some anti-homophily factors

should also be taken into account, especially for controversial topics (Macskassy & Michelson, 2011).

In particular, there are a variety of features that can impact retweetability. For content features, message utility, mention of user handles (Yang et al., 2018), hashtags, URLs (Suh et al., 2010), and emotions (Stieglitz & Dang-Xuan, 2013) often lead to more retweets. For user features, the user’s interest in the topic, opinions, and perceived relevance to their followers/communities can affect the retweeting behavior (Boehmer & Tandoc, 2015; Hoang & Mothe, 2018). From the network perspective, studies based on information propagation models suggest the importance of network topology and parameters in information diffusion (Kumar & Sinha, 2021). While big influencers and stronger ties are more influential individually, it is evident that the more abundant weak ties may play a more dominant role in the propagation of novel information (Bakshy et al., 2012). Other factors such as timing and exposure also affect retweeting (Yoo et al., 2016).

Via retweeting, social media platforms can serve as an effective channel for disseminating useful information and benefiting the public, especially during crises and disasters (Yoo et al., 2016). However, in the meantime, they have also become the hotbed for sharing misinformation. It is a common belief that people fall for fake news due to politically motivated reasoning, but recent evidence contradicts this belief and suggests a lot more than that (Pennycook & Rand, 2021). Lin et al. (2021) conducted an internet survey and attributed users’ motivations to retweet misinformation to socializing, information seeking, or status seeking, not that different from why their share other information online. Wang et al. (2021) studied the spread of misinformation on social media from the perspective of how people process information. According to the Heuristic-Systematic Model (Chaiken, 1980), people process information using two modes: heuristic processing and systematic processing. On one hand, the systematic processing involves cognitive efforts of understanding the content of a message, which may lead to different retweetability of misinformation across topics. On the other hand, the heuristic processing relies more on mental shortcuts and rules of thumb, which explains why a political leader’s behaviors may “nudge” people’s sharing of misinformation. Poor discernment of falsehoods is linked to a lack of careful reasoning and relevant knowledge, as well as to the use of familiarity and source heuristics. Sharing does not necessarily indicate belief (Pennycook et al., 2021). Although most people know that it is important to share accurate news, they may still share misinformation not purposefully but because their attention is distracted from accuracy to other factors (e.g., familiarity, ideology). Websites should implement mechanisms to shift users’ attention to accuracy to increase the quality of information that they share.

Prior studies have investigated the factors that drive the spread of misinformation from three perspectives. First, from the perspective of content and source, misinformation often includes “clickbait” headlines, exaggerated language, and graphic images to attract attention (Baptista & Gradim, 2020). Second, from a user perspective, individuals of older ages or with lower levels of education are more likely to be victims of misinformation (Allen et al., 2020; Georgiou et al., 2020; Grinberg et al., 2019; Xiong & Zuo, 2019), and further share erroneous information (Bessi et al., 2015). Third, from a network perspective, many researchers have adopted information propagation models and epidemic models to characterize the diffusion of misinformation (Allcott & Gentzkow, 2017; Cinelli et al., 2020; Garrett, 2019; Grinberg et al., 2019; Mosleh et al., 2020; Tambuscio et al., 2015; Wood, 2018). On different social media platforms, rumors amplify at various rates, driven by the interaction paradigm imposed by the platform and by the specific communication pattern among users engaged with the topic. Social media networks tend to form “echo chambers” of ideological segregation, which cloud people’s judgment on what to believe or share.

2.3 Research Gaps

Fighting misinformation on the internet is a major challenge. Thanks to the previous works, we have become more capable of detecting misinformation and preventing its dissemination. Even though we have fact-checkers like PolitiFact to verify the information and tell truth from the lies, their fact-checking efforts are by nature reactive and often too late to reverse the damage already caused by fake news. Hence, an even more challenging task in the context of fighting misinformation is to disseminate the fact-checks, i.e., the truth, to the public as fast as possible and as wide as possible. Zhang et al. (2019) built a mathematical model of rumor propagation and found that the initial number of the true information spreaders affects the peak value of the rumor spreaders and the duration of the rumor. Hence, the key to mitigating the damage of rumors is to have more people spread the truth.

Although many works have studied how to detect misinformation and prevent it from spreading, there has been limited research that investigates the spread of the truth. In the race against rumors and lies, the truth is often at a disadvantage. By analyzing 126,000 rumors on Twitter, Vosoughi et al. (2018) found that fake news can diffuse much faster and reach more people than the truth. They attributed this difference to the degree of novelty and the emotional reactions of recipients. In prior literature, believing and spreading misinformation can be attributed to a variety of factors, from an information perspective (e.g., information content and source) (Wang et al., 2021), a user perspective (e.g., individual’s ability and motivation to spot falsehoods) (Allen et al.,

2020; Pennycook et al., 2021), and a network perspective (e.g., group-level and societal factors) (Allcott & Gentzkow, 2017; Scheufele & Krause, 2019). However, the main factors that can facilitate or hinder the spread of fact-checks on the internet remain research gaps and need in-depth investigation.

Swamped and misled by the enormous amount of misinformation on social media, people are in urgent need of more effective and efficient ways to access the truth and hopefully further spread the truth to others. Once fact-checks are carried out and published by valid sources, they need to be quickly disseminated over the internet to put out the wild-fire of misinformation. In this research, we aim to uncover the key factors that affect the dissemination of fact-checks and promote the sharing of fact-checks on social media to mitigate the damage from misinformation.

3 Method

3.1 Data Collection

To investigate the spread of fact-checks on social media, we collected a dataset of statements verified by a well-known fact-checking website, *PolitiFact.com*. PolitiFact is a website that focuses on fact-checking U.S. politics and investigates the accuracy of statements claimed by political figures and viral stories on social media. PolitiFact was awarded the Pulitzer Prize for National Reporting in 2009 for “its fact-checking initiative during the 2008 presidential campaign.” According to PolitiFact’s website, they use “on-the-record interviews and publish a list of sources with every fact-check” with an emphasis on “primary sources and original documentation.” Their fact-checking process includes the following: “a review of what other fact-checkers have found previously; a thorough Google search; a search of online databases; consultation with a variety of experts; a review of publications and a final overall review of available evidence.” PolitiFact rates statements on its trademarked Truth-O-Meter, including six ratings in descending order of truthfulness:

- *True* – The statement is accurate and there is nothing significant missing.
- *Mostly True* – The statement is accurate but needs clarification or additional information.
- *Half True* – The statement is partially accurate but leaves out important details or takes things out of context.
- *Mostly False* – The statement contains an element of truth but ignores critical facts that would give a different impression.
- *False* – The statement is not accurate.
- *Pants on Fire* – The statement is not accurate and makes a ridiculous claim.

From PolitiFact’s website, we scraped a total of 1,003 fact-checks by PolitiFact during six months between November 2020 and May 2021. For each fact-check, we collected the following data: original statement, date of statement, speaker (person or source), truthfulness rating, and tags (a list of keywords assigned by PolitiFact).

PolitiFact keeps track of the sources of all statements. For political figures with a large number of rated statements, the website aggregates all their ratings in a “scorecard,” which summarizes the distribution of the six ratings of truthfulness. The scorecard can say a lot about one’s credibility. From each source’s page, we scraped data such as name, title, description, and scoreboard (i.e., the number of statements in each of the six ratings). As PolitiFact does not show/reveal the view count of fact-check pages on its website, we had to find a second data source to measure the spread/influence of the fact-checks by PolitiFact. Therefore, we looked at PolitiFact’s own Twitter page (<https://twitter.com/PolitiFact>), where it shared its fact-checks and commentaries on recent news stories and statements. Using Twitter API, we collected PolitiFact’s 3,200 statuses (tweets) from its timeline in the same 6-month period (November 2020 ~ May 2021), including tweet ID, date, text, retweet count, and favorite count. These 3,200 statuses include tweets that link to fact-check pages on PolitiFact.com. By matching the URL of fact-check pages, we joined the two datasets together. It is worth noting that PolitiFact does not tweet all its fact-checks on Twitter and some fact-checks may be (re)tweeted multiple times. Hence, the integrated dataset contains a total of 635 unique fact-checks. Table 1 summarizes the distribution of ratings of the 635 fact-checks. Notably, among the statements fact-checked and tweeted by PolitiFact, only 3.6% were rated *True*, whereas the majority were rated *False* (46.6%) or *Pants on Fire* (21.7%). PolitiFact does not choose which fact-checks to share on Twitter based on their truthfulness ratings as demonstrated by the similar distributions between the complete set of 1,003 fact-checks and the 635 tweeted fact-checks in Table 1.

Table 1 Distribution of statements with different ratings

Rating	Fact-checks (count; %)		Tweeted Fact-checks (count; %)	
<i>True</i>	51	5.1%	23	3.6%
<i>Mostly True</i>	49	4.9%	30	4.7%
<i>Half True</i>	83	8.3%	45	7.1%
<i>Mostly False</i>	144	14.4%	103	16.2%
<i>False</i>	468	46.7%	296	46.6%
<i>Pants on Fire</i>	208	20.7%	138	21.7%
Total	1,003	100%	635	100%

3.2 Variables

3.2.1 Dependent Variable

In this research, we investigate different factors that affect the spread of fact-checks on social media. To measure the spread of a fact-check, we use the number of retweets for all tweet(s) containing a link to it. There are cases when PolitiFact’s Twitter account tweeted about the same fact-check multiple times, which may be across different days, weeks, or even months. Therefore, we sum up the retweet counts of all statuses that include a link to the same fact-check page as a *total retweet count*, which is used as the dependent variable in our models.

3.2.2 Independent Variables

Related studies have shown that various factors may affect the spread of information (Allcott & Gentzkow, 2017; Lee et al., 2015). In this research, our dataset includes fact-checks on statements rated and posted by the fact-checking organization, PolitiFact. According to the Heuristic-Systematic Model (Chaiken, 1980), people process information using two modes, namely heuristic processing and systematic processing. To analyze the spread of misinformation, Wang et al. (2021) considered the content of a message, particularly topics extracted by LDA, as the factor of systematic processing. Since our research focuses on the spread of fact-checks on information, the rating of a fact-check should be considered an aspect of systemic/cognitive processing. We focus on the following two factors as independent variables:

Truthfulness In our dataset, each fact-check is on a statement verified by PolitiFact with a six-level rating of truthfulness, from *True* to *Pants on Fire*. The rating of the fact-check may trigger different emotional responses in individuals, which may further affect their intention to spread the information (Vosoughi et al., 2018). To uncover the effects of different ratings on information spread, we treat the truthfulness rating as a categorical variable in our models.

Topics PolitiFact’s fact-checks a variety of statements from different figures or sources, generally in the US political field. The topic of statements may have a critical effect on the level of attention and popularity. Hence, the fact-checks of statements on different topics may result in different levels of spread, as well. Following the approach of (Wang et al., 2021), we conduct a topic analysis on all statements to extract the main topics and study how they affect the spread of fact-checks. Topic modeling is a statistical tool for discovering hidden semantic structures from a collection of texts. In particular, we use a widely used topic model,

Latent Dirichlet Analysis (LDA), to identify the main topics in the collection of statements (Blei et al., 2003). Based on the intuition that documents cover a small number of topics and that topics use a certain set of words, LDA can estimate the document-topic and topic-word distributions from a body of text. For the dataset of 635 statements (November 2020~May 2021), the LDA model identifies two main topics (see Table 2). In particular, Topic 1 involves the development of the COVID-19 pandemic and vaccines, whereas Topic 2 involves a variety of claims related to the 2020 US Presidential Election.

3.2.3 Control Variables

When users encounter fact-checks of stories, the systematic processing of information deals with the contents and ratings, whereas the heuristic processing involves other factors (e.g., characteristics of sources) that may affect users' understanding and sharing of information (Wang et al., 2021). In our model, we control for source-related variables while investigating the main factors of interest.

PolitiFact maintains a profile page for each political figure or group including information such as title, party, and a short description. For those with a large number of rated statements, the website also provides a summary of different ratings as a “scorecard.” For these individuals or groups, we define a score named *credibility* by calculating the weighted average of the six rating scores received (in Table 3) by the percentage of each rating. For example, if an individual's claims include 75% *True* (score = 1) and 25% *False* (score = -1), then his/her credibility score is $75\% \times 1 + 25\% \times (-1) = 0.5$.

During a crisis such as the pandemic, people are overloaded with an enormous amount of information but without the ability to assess its accuracy (Li et al., 2020; Rathore & Farooq, 2020; Van Bavel et al., 2020). Sources such as political leaders and celebrities tend to make louder voices for their established following. On platforms like Twitter, some key opinion leaders' COVID-related claims can quickly become viral (Rufai & Bunce, 2020) and influence people's judgment on the pandemic (Van Bavel et al., 2020; Wang et al., 2021). For our data set,

Table 3 A mapping table to convert ratings to scores

PolitiFact	
Rating	Score
<i>True</i>	1
<i>Mostly True</i>	0.5
<i>Half True</i>	0
<i>Mostly False</i>	-0.5
<i>False</i>	-1
<i>Pants on Fire</i>	-2

PolitiFact claims that, in choosing which statements to fact-check, they take into account factors such as verifiability, significance, and likelihood of propagation. The sources being fact-checked range from democrats and republicans, to individuals, organizations, and platforms. Nevertheless, sources still vary in the number of statements being fact-checked. It is no surprise that the “usual suspects” include the party that holds power, political figures who repeatedly make misleading statements, and popular social media platforms (e.g., Facebook posts and viral images). Hence, for each source, we count the total number of statements that have been fact-checked by PolitiFact and control for this variable, *source statement count*, in our models.

In such a politically divided atmosphere as the current U.S., one's stance in the political spectrum from the “left” (liberal) to the “right” (conservative) may also affect the social influence. By reading each source's profile description on PolitiFact and cross-referencing other online information, we determine its *political leaning* as one of three values: *liberal*, *conservative*, or *NA*. *NA* refers to social media platforms, organizations, and individuals that do not demonstrate a clear political bias.

Since the dependent variable is the total retweet count of tweets linking to a fact-check's page, pages that were tweeted by PolitiFact multiple times and/or spanning a long period gain more exposure. Bakshy et al. (2012) demonstrated the positive effects of exposure to signals on information diffusion. Hence, to control for these factors, we define a variable *tweet count* as the total number of times that a fact-check was tweeted by PolitiFact; we also find the timestamps of the first and the last tweet for each fact-check in the dataset and calculate their difference as a

Table 2 Two topics identified using the LDA model

Topic	Top 20 keywords	Interpretation
1	Biden, vaccine, Joe, covid, vote, Trump, year, people, show, president, Texas, house, elect, win, photo, state illegal, white, border, million	COVID-19 and vaccines
2	Elect, vote, ballot, state, Biden, show, covid, Trump, photo, people, vaccine, president, Joe, voter, video, Georgia, Donald, day, new, capitol	2020 Presidential Election

control variable *time range* (in days). For fact-checks that were only tweeted once, the first tweet is also the last and so the time range is zero.

3.2.4 Descriptive Statistics of Variables

Table 4 summarizes all the variables defined above. Compared to the 8.3% statements out of the 635 that are rated *True* or *Mostly True*, more than two-thirds of all statements are rated either *False* or *Pants on Fire*, as fact-checking websites such as PolitiFact are generally more concerned with debunking false or misleading statements that cause more harm than good. The majority of the sources are neither liberal nor conservative as information spreads quickly and widely through Facebook posts and viral images that do not have a clear political affiliation. The four variables, including *Total Retweet Count*, *Source Statement Count*, *Tweet Count*, and *Time Range*, have distributions that are far from normal. A logarithmic transformation of *Total Retweet Count* normalizes the distribution as shown in Fig. 1. We also transformed *Source Statement Count*, *Tweet Count*, and *Time Range* to their respective logarithmic forms to stabilize these variables for later computations.

The Pearson correlation coefficients in Table 5 demonstrate that numerical variables in the dataset are mostly uncorrelated with the exception between *Source Credibility* and $\log(\text{Source Statement Count})$, between *Source Credibility* and $\log(\text{Time Range})$, and between $\log(\text{Tweet Count})$ and $\log(\text{Time Range})$. In this dataset, influential sources that make more claims and tweets with a shorter *Time Range* generally have lower credibility than other sources, but the absolute correlation coefficients are not high enough to lead to unstable model parameter estimates. Because *Tweet Count*

and *Time Range* are moderately positively correlated on the logarithmic scale, the variance inflation factor (VIF) is used to assess the level of multicollinearity in Section 3.3.

3.3 Models

We employ negative binomial regression to model the effects of the truthfulness ratings, topics, and various source features of the statements on the total retweets of the statements. Negative binomial regression models have been frequently used to capture user engagement on social media including the number of likes, comments, and shares (Bakhshi et al., 2014; He et al., 2018; Lee et al., 2015; Shirish et al., 2021). Compared to the standard Poisson regression model, the negative binomial model is more appropriate as it does not require the variance to be equal to the mean and solves the overdispersion problem that is evident when the Poisson model is applied to our data.

In this work, we express the negative binomial regression model as follows:

$$\begin{aligned} \log(\text{TotalRetweetCount}) = & \beta_0 + \beta_1 \text{True} \\ & + \beta_2 \text{MostlyTrue} + \beta_3 \text{MostlyFalse} \\ & + \beta_4 \text{False} + \beta_5 \text{PantsOnFire} \\ & + \beta_6 \text{TopicCovid} + \beta_7 \text{SourceCredibility} \\ & + \beta_8 \log(\text{SourceStatementCount}) \\ & + \beta_9 \text{Liberal} + \beta_{10} \text{Conservative} \\ & + \beta_{11} \log(\text{TweetCount}) + \beta_{12} \log(\text{TimeRange}) \\ & + \epsilon, \end{aligned} \tag{1}$$

Table 4 Summary of variables collected from 635 statements from November 2020 to May 2021

Variables	Mean	Std. Dev.	Minimum	Maximum	Count
<i>Total Retweet Count</i>	127.8551	274.8686	1	3406	-
<i>Rating – True</i>	-	-	-	-	23 (3.6%)
<i>Rating – Mostly True</i>	-	-	-	-	30 (4.7%)
<i>Rating – Half True</i>	-	-	-	-	45 (7.1%)
<i>Rating – Mostly False</i>	-	-	-	-	103 (16.2%)
<i>Rating – False</i>	-	-	-	-	296 (46.6%)
<i>Rating – Pants on Fire</i>	-	-	-	-	138 (21.7%)
<i>TopicCovid</i>	0.4951	0.2368	0.1454	0.8340	-
<i>TopicElection</i>	0.5049	0.2368	0.1660	0.8546	-
<i>Source Credibility</i>	-0.8400	0.4859	-2	1	-
<i>Source Statement Count</i>	614.7228	573.6298	1	1300	-
<i>Source – Liberal</i>	-	-	-	-	70 (11.0%)
<i>Source – Conservative</i>	-	-	-	-	191 (30.1%)
<i>Source – NA</i>	-	-	-	-	374 (58.9%)
<i>Tweet Count</i>	1.9921	1.1240	1	10	-
<i>Time Range</i>	4.0252	13.2284	0	131	-

Fig. 1 Histograms of *Total Retweet Count*, *Source Statement Count*, *Tweet Count*, *Time Range*, and their log transforms

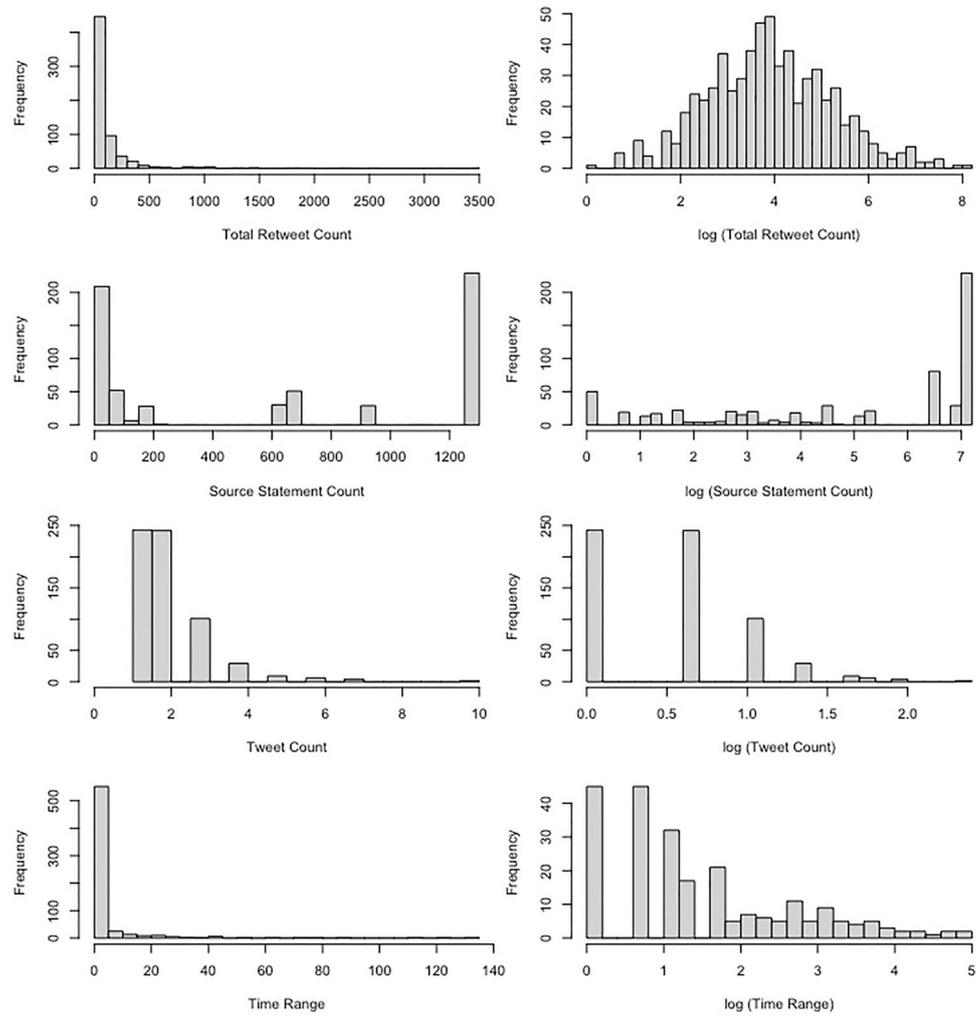


Table 5 Pairwise Pearson correlation coefficients among the numerical variables

Variables	$\log(\text{Total Retweet Count})$	<i>TopicCovid</i>	<i>Source Credibility</i>	$\log(\text{Source Statement Count})$	$\log(\text{Tweet Count})$	$\log(\text{Time Range})$
$\log(\text{Total Retweet Count})$	1.000	-0.017	-0.164	0.047	0.602	0.404
<i>TopicCovid</i>	-	1.000	-0.028	-0.020	0.013	-0.013
<i>Source Credibility</i>	-	-	1.000	-0.438***	0.057	0.101**
$\log(\text{Source Statement Count})$	-	-	-	1.000	-0.010	-0.039
$\log(\text{Tweet Count})$	-	-	-	-	1.000	0.708***
$\log(\text{Time Range})$	-	-	-	-	-	1.000

Pearson correlation coefficients that are significant at the 0.1%, 1%, and 5% significance levels are marked by ***, **, and *, respectively

where the dependent variable measures the total number of retweets of each statement. For the first factor, *Truthfulness*, we use five dummy variables to indicate whether the statement receives a specific rating, while the rating of *HalfTrue* is set as the baseline level. *TopicCovid* and *TopicElection* are numbers between 0 and 1 measuring the statement’s likelihood of revolving around the COVID-19

and vaccines (Topic 1) and the 2020 Presidential Election (Topic 2), respectively. For each statement, the sum of its *TopicCovid* and *TopicElection* is always 1, so only *TopicCovid* is included in the model. Based on the source of each statement, we calculated *Source Credibility* using Table 3 and transformed the total number of statements made by the same source using the logarithmic transformation to

take care of the skewness in the data. Two binary variables, *Liberal* and *Conservative*, are used to capture the political bias of the source; the baseline level is *NA* which represents a source without a clear political leaning. *Tweet Count* controls for the number of times a fact-check page has been tweeted by PolitiFact while *Time Range*, controls for the number of days difference between the very first tweet and the very last tweet. The VIF results show that all independent variables have VIF values less than 5, indicating multicollinearity is not a problem in our model. The parameters in the negative binomial regression model are estimated using an alternating iteration process until convergence is reached (McCullagh & Nelder, 1989).

4 Results & Discussion

4.1 Results

We present the negative binomial regression estimates in Table 6. The full regression model defined in Eq. (1) is Full Model, whereas the reduced regression model using only the control variables is Reduced Model. While the coefficient estimates and significance levels of the control variables are consistent across the two models, the Akaike Information Criterion (AIC) and $2 \times \log$ -likelihood indicate that the Full Model performs better than the Reduced Model, implying the usefulness of the independent variables, especially the statement truthfulness rating. A non-zero dispersion

parameter estimate indicates our data are overdispersed and are better fitted using the negative binomial regression model than the standard Poisson model.

The results show that the *truthfulness* rating is a key determinant factor for the attention received by the fact-checked statements as measured by the total retweet count. Compared to statements that are rated *Half True* as the baseline, statements that are either *True* or *False* tend to earn significantly more retweets on average. More specifically, holding all other variables constant, a *True* statement earns $\exp(1.8892) = 6.61$ times (or 561% increase), and a *Pants on Fire* statement earns $\exp(1.6709) = 5.32$ times (or equivalently 432% increase) of the total retweets received by a *Half True* statement on average. To illustrate the nonlinear relationship between the truthfulness ratings and the total retweets, we plot the 95% confidence intervals of the ratios of five statement ratings against the baseline (*Half True*) in Fig. 2. The dashed line at the ratio of 1 represents the baseline of statements rated *Half True*. The centers, i.e., the medians, of these intervals are marked by solid circles. This indicates that people are more inclined to spread true or overly fabricated information through social media than half true information. Fact-checks that confirm true statements tend to receive the most retweets.

From the perspective of PolitiFact’s readers, when they see a statement being fact-checked, a rating that lands on either end of the truthfulness spectrum seems to attract more attention from the public and give them more motivation to pass it along. In contrast, if a statement is rated partially true

Table 6 The estimated effects from the negative binomial regression models

Dependent Variable: Total Retweet Count		Reduced Model		Full Model	
		Estimate	(s.e.)	Estimate	(s.e.)
Independent Variables	<i>Rating – True</i>			1.8892	(0.2449) ***
	<i>Rating – Mostly True</i>			0.9328	(0.2239) ***
	<i>Rating – Mostly False</i>			0.9268	(0.1771) ***
	<i>Rating – False</i>			1.4561	(0.1690) ***
	<i>Rating – Pants on Fire</i>			1.6709	(0.1842) ***
	<i>TopicCovid</i>			-0.1412	(0.1567)
Control Variables	<i>Source Credibility</i>	-0.5966	(0.1222) ***	-0.4837	(0.1308) ***
	\log (<i>Source Statement Count</i>)	0.0137	(0.0222)	-0.0003	(0.0211)
	<i>Source – Liberal</i>	0.2017	(0.2064)	0.2273	(0.1957)
	<i>Source – Conservative</i>	0.4430	(0.1390) **	0.4367	(0.1331) **
	\log (<i>Tweet Count</i>)	1.5068	(0.1140) ***	1.4238	(0.1075) ***
	\log (<i>Time Range</i>)	0.0079	(0.0107)	0.0093	(0.0101)
	Constant	2.9551	(0.2216) ***	1.8602	(0.2563) ***
Others	Dispersion Parameter	1.0346	(0.0524)	1.1816	(0.0608)
	AIC	6982.2		6890.3	
	$2 \times \log$ -likelihood	-6966.2		-6862.3	

Estimates that are significant at the 0.1%, 1%, and 5% significance levels are marked by ***, **, and *, respectively. A statement that is rated as *Half True* is used as the baseline. A source that is neither liberal nor conservative is used as the baseline

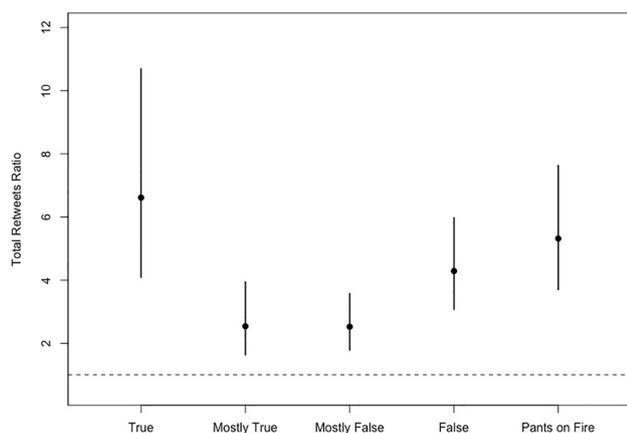


Fig. 2 The 95% confidence intervals of the ratios of total retweets received by *True*, *Mostly True*, *Mostly False*, *False*, and *Pants on Fire* statements, compared to that of *Half True* statements. The centers of the 95% confidence intervals are marked by solid circles

or false, readers have the impression of uncertainty and are less motivated to spread the information. Given this bimodal distribution of total retweet count, we cannot help but think of the notorious J-shaped distribution often revealed in product reviews with many 5-star ratings, some 1-star ratings, and hardly any ratings in between. Hu et al. (2007) attribute the J-shaped distribution to two biases: “(1) purchasing bias - only consumers with a favorable disposition towards a product purchase the product and have the opportunity to write a product review, and (2) under-reporting bias - consumers with polarized (either positive or negative) reviews are more likely to report their reviews than consumers with moderate reviews.” Our results show that similar biases seem to apply when people process and react to information of different truthfulness on the internet: (1) “following bias” - readers/followers of fact-checkers like PolitiFact tend to be people who are more passionate in debunking fake news and are more sensitive to the ruling of misinformation; and (2) “under-reporting bias” - people with more polarized views are more likely to spread information than those with moderate views.

In our analysis, the topic of a statement does not affect the total retweets of the statement regardless of whether this statement is more focused on the COVID-19 situation or the 2020 Presidential Election. Fact-checkers such as PolitiFact are set to collect, investigate, and verify the truthfulness of statements that focused on a variety of topics. Statements on any topic could turn out to be true or false. COVID-19 and the 2020 Presidential Election happened to be the main issues during the period of our data collection. The majority of statements in our dataset are closely related to one of the two topics or both. At least in our dataset, there is no evidence of either topic being a strong factor that affects the spread of fact-checks.

Among the control variables, the credibility of the source has a significantly negative impact on the total retweet count while a conservative source and tweet count have positive effects on the total retweet count. A statement from a source with a credibility score of 1 (that is most of its statements are rated as *True*) gains only 38.01% (between 29.42% and 49.11% at a 95% confidence level) of the retweets of a statement from a source with a credibility score of -1 (generally rated as *False*), holding all other factors the same. Less credible sources tend to make more false or sometimes even ridiculous statements, which often attract more attention and lead to more retweets. While there is no significant difference in total retweet count between the liberal sources and the sources demonstrating no clear political leaning, the conservative sources tend to collect 54.76% more retweets (between 19.24% and 100.87% at a 95% confidence level) than those with no clear political bias. Twitter users are more likely to spread fact-checks on statements from conservative sources. When PolitiFact doubles the number of tweets made for fact-check pages, the total retweet count would increase by 168.28% (between 131.84% and 210.46% at a 95% confidence level). This is not surprising because statements that were tweeted by PolitiFact multiple times tend to be the ones that are concerned with issues/events gaining wide popularity and concerns. More exposure to the statements and their ratings tends to accumulate more retweets over time, suggesting the critical impact made by fact-checking organizations like PolitiFact on spreading information.

Other control variables do not appear to play any significant role in affecting the total retweet count. For example, although individuals/sources who contribute more statements fact-checked by PolitiFact may have been the usual suspects of misinformation, what they had to say, true or false, did not spark more or less attention than those less vocal/influential sources. Holding all other factors constant, a statement that has been tweeted on social media for a long period of time receives about the same number of total retweets as a statement that has only been recently tweeted by PolitiFact.

4.2 Discussion

Combating misinformation is a challenging task that cannot be simply resolved at the individual level. As misinformation emerges in group-level processes shaped by societal dynamics, we need to take a system’s approach to better understand “the vulnerabilities of individuals, institutions, and society” to misinformation disseminated through ubiquitous online channels (Lazer et al., 2018; Scheufele & Krause, 2019). From the findings of our research, we can gain some new insights into how fact-checks of stories are spread on social media and how we can effectively disseminate the truth.

As people read fact-checks, they make cognitive efforts to systematically process the information, particularly the content in question and the judgment. As shown in our models, the truthfulness rating on a statement is a key factor that affects individuals' tendency to share fact-checks with others. Noticeably, this relationship is nonlinear. As compared to a *Half True* rating, a more conclusive rating, either at the *True* or the *False* end of the truthfulness spectrum, tend to get more retweets. This finding can be attributed to confirmation bias (Modgil et al., 2021) and under-reporting bias (Hu et al., 2007) in that people tend to seek and favor information that supports one's prior beliefs and, in response, to disseminate the information to others. Fact-checks with a conclusive rating, either true or false, make it a lot easier for people to spot information confirming their beliefs and hence tend to be shared more. On the one hand, it is certainly good to see that the confirmation of true statements and the debunking of false statements get spread further and help to set the record straight. On the other hand, it is worth keeping in mind that stories can be misleading when words are taken out of context or certain aspects are omitted in some stories/statements. Truthfulness is a spectrum on which each statement being fact-checked lands. These fact-checks, with ratings such as *Mostly True*, *Half True*, or *Mostly False*, provide a comprehensive analysis of the story from multiple perspectives and with supporting/opposing evidence. They deserve just equal, if not more, attention than those with a more conclusive rating. It is not to say that users should be blamed for not retweeting enough. After all, it is human nature to be more attracted to and responsive to words that choose/stand a clear side. Perhaps, to spread impartial and objective fact-checks, there is more work to be done with those that are not completely true or false. For a statement that is rated *Half True*, rather than only showing the rating as a thumbnail image, perhaps highlighting the words/phrases that make it partially false would make the fact-check noticeable and compelling. For important fact-checks of statements that call for more attention, e.g., claims by major political leaders or those potentially causing wide confusion/misunderstanding, they may deserve to be shared more than once in media platforms (e.g., Twitter) to increase their coverage and impact. Rumors and misinformation can be spread over the internet so quickly and widely. Users should be encouraged to pay attention to fact-checks of statements that are partially true/false as much as those with a definite ruling. Furthermore, by sharing more fact-checks with others, users can help effectively prevent the spread of misinformation on the internet and mitigate the damage.

According to (Wang et al., 2021), different topics of COVID-19-related misinformation vary in their popularity, with conspiracy theories being retweeted the most. When it comes to the spreading of fact-checks, our model does not show the topic as a significant factor that affects retweetability.

Nevertheless, it is worth noting that our dataset covers a unique period in 2020 when the U.S. Presidential Election and the COVID-19 pandemic are two predominant topics. Our topic model only focuses on these two topics and calculates to what extent each fact-check is related to them, respectively. Although there is no significant difference between the two main topics, this does not necessarily mean that all sub-topics, if further divided, are equally popular or retweetable. With that said, the lack of variation in retweetability for fact-checks across different topics means that any fact-checks, even those on less eye-catching stories, with proper distribution and promotion, can obtain a substantial level of dissemination, as they all deserve.

In our models, we control for source-related variables that may affect users' heuristic processing of information (Wang et al., 2021). The results show that, among these control variables, source credibility, political leaning, and tweet count can significantly impact the spread of fact-checks. Fact-checks of stories from sources with lower credibility, e.g., Facebook posts and viral images, seem to gain more attention and retweets. For political leaning, fact-checks of statements originating from conservative sources tend to attract more public attention. In addition, when a fact-check of a statement is posted multiple times by PolitiFact on social media, it is likely to collect more retweets and gather more attention. This confirms the positive effect of exposure on information (Bakshy et al., 2012). Therefore, to increase the spread and impact of a fact-check, a simple method would be to post it multiple times, perhaps with follow-up investigation and analysis as the story develops.

Pennycook et al. (2020) suggest that there is a disconnection between what people believe and what they share on social media. Although most people are for sharing accurate information, many may become the spreaders of rumors on social media. Their sharing of misinformation is not necessarily due to misinformed beliefs, but a result of thoughtlessness or negligence. Whether it is their intention or not, their sharing facilitates and amplifies the propagation of misinformation. Likewise, sharing of fact-checks may not necessarily require beliefs. Since rumors tend to travel faster than the truth (Vosoughi et al., 2018), to beat the spread of rumors we need as many users as possible to help disseminate the truth, regardless of their beliefs. As long as a fact-check is shared by one more user, it can count as a win for the truth and we are one step closer to catching up on the spread of misinformation. Properly designed training protocols can also enhance the ability of online users to recognize fake news and spread the truth (Soetekouw & Angelopoulos, 2022). Reminding users to focus on accuracy can influence what they share on social media (Pennycook et al., 2021). Fact-checkers such as PolitiFact establish and maintain a trustworthy

reputation by continuously undertaking an unbiased, thorough, and comprehensive fact-checking process (e.g., selection, investigation, and reporting). To better promote the sharing of fact-checks, they should also remind people about their credentials and the rigidity of their fact-checking process.

5 Conclusions & Future Directions

Fake news is not a new phenomenon. However, aided by social media's features such as retweeting and sharing, misinformation can propagate like wildfire. To combat misinformation, this research studies different factors that may affect the spread of fact-checks. Our analytical models identify the truthfulness rating as a significant factor: conclusive fact-checks (either *True* or *False*) get shared more. Furthermore, the source credibility and the time length of sharing also affect the spread of fact-checks over the internet.

To counter the dissemination of misinformation is by no means an easy task and it calls for an integrated strategy that combines efforts from multiple sides of our society (Rodrigo et al., 2022). From the individual user side, users should be encouraged to follow a set of tenets that promote the truth and dismiss the lies. Promoting media literacy and educating the public on digital resilience can enhance awareness around source credibility and detection of misinformation. Meanwhile, from the business side, rather than chasing high engagement rates and letting misinformation unhinged (Modgil et al., 2021), social media platforms must take action to hinder the spread of misinformation by enforcing fact-checking standards and promoting the sharing of truthful information. It may be unrealistic to expect technology companies to police every single piece of content posted on their platforms, especially when doing so might hurt them, thus, the governmental agencies should actively work with these platforms by providing guidelines and adjudication to combat fictitious and malicious content on the internet. The battle against fake news to create a better internet infrastructure will not be successful without the combined efforts of the users, the platforms, and the governments.

In the future, we suggest extending this research in the following directions. First, rather than only studying fact-checks by one organization (PolitiFact) highly focused on U.S. politics, we will expand the dataset by analyzing larger collections of fact-checks on a wider range of topics by multiple fact-checkers (with different credentials). Second, we will investigate other factors that may affect the spread of truth and misinformation, such as the lapse between the emerging and fact-checking of fake news, the credibility and

biases of fact-checkers, etc. Last, we will collect the content and sentiments in users' comments on fake news and the corresponding fact-checks, which may provide insights into how/why people respond to information and share it.

Declarations

Conflict of Interest The authors did not receive support from any organization for the manuscript. The authors have no relevant financial or non-financial interests to disclose.

References

- Abdullah, N. A., Nishioka, D., Tanaka, Y., & Murayama, Y. (2017). Why I retweet? Exploring user's perspective on decision-making of information spreading during disasters. *Proceedings of the Annual Hawaii International Conference on System Sciences, 2017-January*, 432–441. <https://doi.org/10.24251/HICSS.2017.053>
- Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/JEP.31.2.211>
- Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. <https://doi.org/10.1126/SCIADV.AAY3539>
- Au, C. H., Ho, K. K. W., & Chiu, D. K. W. (2021). The role of online misinformation and fake news in ideological polarization: Barriers, catalysts, and implications. *Information Systems Frontiers*, 1–24. <https://doi.org/10.1007/s10796-021-10133-9>
- Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014). Faces engage us: Photos with faces attract more likes and comments on instagram. *Conference on Human Factors in Computing Systems - Proceedings*, 965–974. <https://doi.org/10.1145/2556288.2557403>
- Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web*, 519–528. <https://doi.org/10.1145/2187836.2187907>
- Baptista, J. P., & Gradim, A. (2020). Understanding fake news consumption: A review. *Social Sciences* 2020, 9(10), 185. <https://doi.org/10.3390/SOCSCI9100185>
- Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Science vs conspiracy: Collective narratives in the age of misinformation. *PLoS One*, 10(2), e0118093. <https://doi.org/10.1371/JOURNAL.PONE.0118093>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993–1022. <https://doi.org/10.1162/jmlr.2003.3.4-5.993>
- Boehmer, J., & Tandoc, E. C. (2015). Why we retweet: Factors influencing intentions to share sport news on Twitter. *International Journal of Sport Communication*, 8(2), 212–232. <https://doi.org/10.1123/IJSC.2015-0011>
- Boyd, D., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. *Proceedings of the Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2010.412>

- Carlo, G., Eisenberg, N., Troyer, D., Switzer, G., & Speer, A. L. (1991). The altruistic personality: In what contexts is it apparent? *Journal of Personality and Social Psychology*, 61(3), 450–458. <https://doi.org/10.1037/0022-3514.61.3.450>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>
- Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-based helping: Is it selflessly or selfishly motivated? *Journal of Personality and Social Psychology*, 52(4), 749–758. <https://doi.org/10.1037/0022-3514.52.4.749>
- Cinelli, M., Quattrocioni, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*, 10(1), 16598. <https://doi.org/10.1038/s41598-020-73510-5>
- Conover, M. D., Gonçalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F. (2011). Predicting the political alignment of twitter users. *Proceedings – 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011*, 192–199. <https://doi.org/10.1109/PASSAT/SocialCom.2011.34>
- Dovidio, J. F. (1984). Helping behavior and altruism: An empirical and conceptual overview. *Advances in Experimental Social Psychology*, 17(C), 361–427. [https://doi.org/10.1016/S0065-2601\(08\)60123-9](https://doi.org/10.1016/S0065-2601(08)60123-9)
- Garrett, R. K. (2019). Social media’s contribution to political misperceptions in U.S. Presidential elections. *PLoS One*, 14(3), e0213500. <https://doi.org/10.1371/JOURNAL.PONE.0213500>
- Georgiou, N., Delfabbro, P., & Balzan, R. (2020). COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. *Personality and Individual Differences*, 166, 110201. <https://doi.org/10.1016/J.PAID.2020.110201>
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/SCIENCE.AAU2706>
- He, S., Rui, H., & Whinston, A. B. (2018). Social media strategies in product-harm crises. *Information Systems Research*, 29(2), 362–380. <https://doi.org/10.1287/isre.2017.0707>
- Hoang, T. B. N., & Mothe, J. (2018). Predicting information diffusion on Twitter – Analysis of predictive features. *Journal of Computational Science*, 28, 257–264. <https://doi.org/10.1016/J.JOCS.2017.10.010>
- Hopf, H., Krief, A., Mehta, G., & Matlin, S. A. (2019). Fake science and the knowledge crisis: Ignorance can be fatal. In *Royal Society Open Science* (Vol.6, Issue 5). <https://doi.org/10.1098/rsos.190161>
- Hu, N., Pavlou, P. A., Zhang, J., & Jennifer. (2007). Why do online product reviews have a J-shaped distribution? Overcoming biases in online word-of-mouth communication. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.2380298>
- Hwang, Y., & Shim, H. (2010). Opinion leadership on Twitter and Twitter use: Motivations and patterns of Twitter use and case study of opinion leaders on Twitter. *Korean Journal of Broadcasting*, 24(6), 365–404.
- Kumar, K. P. K., & Geethakumari, G. (2014). Detecting misinformation in online social networks using cognitive psychology. *Human-Centric Computing and Information Sciences*, 2014 4:1(1), 1–22. <https://doi.org/10.1186/S13673-014-0014-X>
- Kumar, P., & Sinha, A. (2021). Information diffusion modeling and analysis for socially interacting networks. *Social Network Analysis and Mining*, 11, 11. <https://doi.org/10.1007/s13278-020-00719-7>
- Kumar, S., & Shah, N. (2018). *False Information on Web and Social Media: A Survey*. Retrieved April 19, 2019, from <https://arxiv.org/abs/1804.08559>
- Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, 18, 18–66. https://doi.org/10.1111/j.1467-8705.2012.02056_3.x
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., & Zittrain, J. L. (2018). The science of fake news: Addressing fake news requires a multidisciplinary effort. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lee, J., Agrawal, M., & Rao, H. R. (2015). Message diffusion through social network service: The case of rumor and non-rumor related tweets during Boston bombing 2013. *Information Systems Frontiers*, 17(5), 997–1005. <https://doi.org/10.1007/s10796-015-9568-z>
- Lee, M., Kim, H., & Kim, O. (2015). Why do people retweet a tweet?: Altruistic, egoistic, and reciprocity motivations for retweeting. *Psychologia*, 58, 189–201.
- Li, Y., Twersky, S., Ignace, K., Zhao, M., Purandare, R., Bennett-Jones, B., & Weaver, S. R. (2020). Constructing and communicating COVID-19 stigma on twitter: A content analysis of tweets during the early stage of the COVID-19 outbreak. *International Journal of Environmental Research and Public Health*, 17(18), 1–12. <https://doi.org/10.3390/ijerph17186847>
- Lin, T. C., Huang, S. L., & Liao, W. X. (2021). Examining the antecedents of everyday rumor retransmission. *Information Technology & People*. <https://doi.org/10.1108/ITP-09-2020-0667>
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12), 3412–3427. <https://doi.org/10.1287/mnsc.2015.2304>
- Mackassay, S., & Michelson, M. (2011). Why do people retweet? Antihomophily wins the day! *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), 209–216. Retrieved July 21, 2021, from <https://ojs.aaai.org/index.php/ICWSM/article/view/14110>
- McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models* (2nd ed.). Chapman and Hall.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. In *Annual Review of Sociology* (Vol. 27, pp. 415–444). <https://doi.org/10.1146/annurev.soc.27.1.415>
- Modgil, S., Singh, R. K., Gupta, S., & Dennehy, D. (2021). A confirmation bias view on social media induced polarisation during Covid-19. *Information Systems Frontiers*, 1, 1–25. <https://doi.org/10.1007/s10796-021-10222-9>
- Mosleh, M., Pennycook, G., & Rand, D. G. (2020). Self-reported willingness to share political news articles in online surveys correlates with actual sharing on Twitter. *PLoS One*, 15(2), e0228882. <https://doi.org/10.1371/JOURNAL.PONE.0228882>
- Olan, F., Jayawickrama, U., Arakpogun, E. O., Suklan, J., & Liu, S. (2022). Fake news on Social Media: the impact on society. *Information Systems Frontiers*, 1, 1–16. <https://doi.org/10.1007/s10796-022-10242-z>
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce

- misinformation online. *Nature*, 592(7855), 590–595. <https://doi.org/10.1038/s41586-021-03344-2>
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy nudge intervention. *Psychological Science*, 31(7), 770–780. <https://doi.org/10.1177/0956797620939054>
- Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2521–2526. <https://doi.org/10.1073/PNAS.1806781116/-DCSUPPLEMENTAL>
- Pennycook, G., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/J.TICS.2021.02.007>
- Pewresearch (2020). 64% in U.S. say social media have a mostly negative effect on country today | Pew Research Center. Retrieved July 23, 2021, from <https://www.pewresearch.org/fact-tank/2020/10/15/64-of-americans-say-social-media-have-a-mostly-negative-effect-on-the-way-things-are-going-in-the-u-s-today/>
- Rabelo, J., Prudencio, R. B. C., & Barros, F. (2012a). Collective classification for sentiment analysis in social networks. *2012 IEEE 24th International Conference on Tools with Artificial Intelligence*, 958–963. <https://doi.org/10.1109/ICTAI.2012.135>
- Rabelo, J., Prudencio, R. B. C., & Barros, F. (2012b). Using link structure to infer opinions in social networks. *IEEE International Conference on Systems, Man, and Cybernetics*, 681–685.
- Rajadesingan, A., & Liu, H. (2014). Identifying users with opposing opinions in Twitter debates. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8393 LNCS, 153–160. https://doi.org/10.1007/978-3-319-05579-4_19
- Rathore, F. A., & Farooq, F. (2020). Information overload and infodemic in the COVID-19 pandemic. *Journal of the Pakistan Medical Association*, 70(5), S162–S165. <https://doi.org/10.5455/JPMA.38>
- Recuero, R., Araujo, R., & Zago, G. (2011). How does social capital affect retweets? *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), 305–312. Retrieved July 23, 2021, from <https://ojs.aaai.org/index.php/ICWSM/article/view/14115>
- Rodrigo, P., Arakpogun, E. O., Vu, M. C., Olan, F., & Djafarova, E. (2022). Can you be mindful? The effectiveness of mindfulness-driven interventions in enhancing the digital resilience to fake news on COVID-19. *Information Systems Frontiers*, 1, 1–21. <https://doi.org/10.1007/s10796-022-10258-5>
- Rozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G. ... van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10). <https://doi.org/10.1098/R SOS.2011199>
- Rufai, S. R., & Bunce, C. (2020). World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. *Journal of Public Health*, 42(3), 510–516. <https://doi.org/10.1093/PUBMED/FDAA049>
- Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences of the United States of America*, 116(16), 7662–7669. <https://doi.org/10.1073/pnas.1805871115>
- Shirish, A., Srivastava, S. C., & Chandra, S. (2021). Impact of mobile connectivity and freedom on fake news propensity during the COVID-19 pandemic: a cross-country empirical examination. *European Journal of Information Systems*, 30(3), 322–341. <https://doi.org/10.1080/0960085X.2021.1886614>
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter*, 19(1), 22–36. <https://doi.org/10.1145/3137597.3137600>
- Shu, K., Wang, S., & Liu, H. (2019). Beyond news contents: The role of social context for fake news detection. *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, 312–320. <https://doi.org/10.1145/3289600.3290994>
- Sloan, L., Quan-Haase, A., & Rubin, V. L. (2017). Deception detection and rumor debunking for social media. In *The SAGE Handbook of Social Media Research Methods* (pp. 342–363). <https://doi.org/10.4135/9781473983847.n21>
- Soetekouw, L., & Angelopoulos, S. (2022). Digital resilience through training protocols: Learning to identify fake news on social media. *Information Systems Frontiers*, 1, 1–17. <https://doi.org/10.1007/s10796-021-10240-7>
- Speriosu, M., Sudan, N., Upadhyay, S., & Baldrige, J. (2011). Twitter polarity classification with label propagation over lexical links and the follower graph. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 53–56.
- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media - Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217–248. <https://doi.org/10.2753/MIS0742-122290408>
- Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in twitter network. *Proceedings - SocialCom 2010: 2nd IEEE International Conference on Social Computing, PASSAT 2010: 2nd IEEE International Conference on Privacy, Security, Risk and Trust*, 177–184. <https://doi.org/10.1109/SOCIALCOM.2010.33>
- Tambuscio, M., Ruffo, G., Flammini, A., & Menczer, F. (2015). Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. *WWW 2015 Companion - Proceedings of the 24th International Conference on World Wide Web*, 977–982. <https://doi.org/10.1145/2740908.2742572>
- Tan, C., Lee, L., Tang, J., Jiang, L., Zhou, M., & Li, P. (2011). User-level sentiment analysis incorporating social networks. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '11*, 136, 1397. <https://doi.org/10.1145/2020408.2020614>
- Tasnim, S., Hossain, M. M., & Mazumder, H. (2020). Impact of rumors and misinformation on COVID-19 in social media. *Journal of Preventive Medicine and Public Health*, 53(3), 171. <https://doi.org/10.3961/JPMMPH.20.094>
- Tran, T., Valecha, R., Rad, P., & Rao, H. R. (2021). An investigation of misinformation harms related to social media during two humanitarian crises. *Information Systems Frontiers*, 23(4), 931–939. <https://doi.org/10.1007/s10796-020-10088-3>
- Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M. ... Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. In *Nature Human Behaviour* (Vol.4, Issue 5, pp.460–471). Nature Publishing Group. <https://doi.org/10.1038/s41562-020-0884-z>
- Vicario, M., Del Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, 113(3), 554–559. <https://doi.org/10.1073/pnas.1517441113>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/SCIENCE.AAP9559>

- Wang, X., Zhang, M., Fan, W., & Zhao, K. (2021). Understanding the spread of COVID-19 misinformation on social media: The effects of topics and a political leader's nudge. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/ASI.24576>
- Wasko, M. M., & Faraj, S. (2000). It is what one does": Why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems*, 9(2–3), 155–173. [https://doi.org/10.1016/S0963-8687\(00\)00045-7](https://doi.org/10.1016/S0963-8687(00)00045-7)
- Wong, F., Tan, C., Sen, S., & Chiang, M. (2013). Quantifying political leaning from tweets and retweets. *International AAAI Conference on Weblogs and Social Media (ICWSM)*.
- Wood, M. J. (2018). Propagating and debunking conspiracy theories on Twitter during the 2015–2016 Zika Virus Outbreak. *Cyberpsychology Behavior and Social Networking*, 21(8), 485. <https://doi.org/10.1089/CYBER.2017.0669>
- Wu, L., Morstatter, F., Carley, K. M., & Liu, H. (2019). Misinformation in Social Media: Definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter*, 21(2), 80–90. <https://doi.org/10.1145/3373464.3373475>
- Xiong, J., & Zuo, M. (2019). How does family support work when older adults obtain information from mobile internet? *Information Technology and People*, 32(6), 1496–1516. <https://doi.org/10.1108/ITP-02-2018-0060>
- Yang, Q., Tufts, C., Ungar, L., Guntuku, S., & Merchant, R. (2018). To retweet or not to retweet: Understanding what features of cardiovascular tweets influence their retransmission. *Journal of Health Communication*, 23(12), 1026. <https://doi.org/10.1080/10810730.2018.1540671>
- Yoo, E., Rand, W., Eftekhari, M., & Rabinovich, E. (2016). Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises. *Journal of Operations Management*, 45. <https://doi.org/10.1016/j.jom.2016.05.007>
- Yusof, A. N. M., Muuti, M. Z., Ariffin, L. A., & Tan, M. K. M. (2020). Sharing Information on COVID-19: the ethical challenges in the Malaysian setting. *Asian Bioethics Review*, 12(3), 349. <https://doi.org/10.1007/S41649-020-00132-4>
- Zhang, J. P., Guo, H. M., Jing, W. J., & Jin, Z. (2019). Dynamic analysis of rumor propagation model based on true information spreader. *Wuli Xuebao/Acta Physica Sinica*, 68(15), 150501–150501. <https://doi.org/10.7498/aps.68.20190191>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Jiexun Li is an Associate Professor in the Department of Decision Sciences, College of Business & Economics, at Western Washington University. He earned his Ph.D. in MIS from the Eller College of Management at the University of Arizona, M.S. and B.S. in MIS at Tsinghua University in China. His research interests include data mining, business analytics, social media analytics, and health informatics. His research has appeared in journals including *Journal of Management Information Systems*, *Decision Support Systems*, *IEEE Transactions*, *Journal for the American Society of Information Systems and Technology*, *Journal for the Association of Information Systems*, *Bioinformatics*, *Communications of the ACM*, *Information Systems Frontiers*, and many others.

Xiaohui Chang is a Toomey Faculty Fellow and Associate Professor of Business Analytics in the College of Business at Oregon State University. Dr. Chang earned her bachelor's degrees in Economics and Statistics (Honors) from the University of Chicago, and her Ph.D. in Statistics from the University of Chicago. Her research interest includes machine learning, business analytics, spatial statistics, and spatio-temporal modeling. She is experienced in developing novel and efficient methods for important applications in business analytics, information systems, finance, and environmental science. Her research work has been published in the *Journal of the American Statistical Association*, *Journal of Machine Learning Research*, *IEEE Transactions*, *Biometrics*, *Expert Systems with Applications*, *Information Technology & People*, *Quantitative Finance*, and many others.